



Traveled distances: New insights into the role of optic flow

Matteo Mossio^{a,b,*,1}, Manuel Vidal^{a,c,1}, Alain Berthoz^a

^a *Laboratoire de Physiologie de la Perception et de l'Action, Collège de France, Paris, France*

^b *IHPST, CNRS/Université Paris 1 Panthéon-Sorbonne, Paris, France*

^c *Max Planck Institute for Biological Cybernetics, Tübingen, Germany*

Received 11 June 2007; received in revised form 5 November 2007

Abstract

In this study, we addressed four related issues concerning the estimation of traveled distances in a distance-matching visual task, using a virtual reality (VR) setup. Firstly, we found that when explicit counting strategies were blocked by an interfering dual task, the performance of 35% of subjects was strongly impaired. Secondly, we found that, when encoding and test phases took place in similar perceptual contexts, subjects' performance could be extremely accurate, which suggests that the inaccuracy and variability reported in previous studies could stem from the use of inefficient mechanisms to building context-independent representations. Thirdly, by systematically manipulating the visual cues available, we ascertained that depth cues and texture regularity were not necessary to estimate traveled distances accurately. Fourthly, we evidenced two distinct groups of subjects according to their dependence on the invariance of speed. While performance remained accurate in some subjects when we manipulated the speed of the test phase it was severely impaired in other subjects, whose strategy seemed to rely on an implicit, time-based estimation. We suggest that the existence of these different groups could account for the inaccuracy and variability observed in previous studies.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Traveled distances; Optic flow; Cognitive load; Virtual reality

1. General introduction

1.1. Traveled distance and static distance perception

One of the fundamental issues of spatial cognition concerns the mechanisms underlying the perception of distances. Humans can potentially use different sources of information in order to estimate distances: the literature draws a classical distinction between *dynamic* and *static* information, depending on whether or not the information is derived from physical motion along the distance to be estimated (Palmer, 1999; Sun, Campos, Chan, Young, & Ellard, 2004). Studies focusing on the contribution of static informa-

tion interpret mechanisms underlying distance perception as being equivalent to those underlying *depth* perception (Marr, 1982; Palmer, 1999): in this approach, the information contributing to the estimation of distances is mainly visual, although in some cases other sensory modalities, like hearing, may be involved (Loomis, Klatzky, Philbeck, & Golledge, 1998). In an alternative approach, distance perception should rather be seen as an estimation of *traveled* distances, relying on the processing of dynamic information generated by the observer's self-motion (Gibson, 1979). These two approaches may in fact be complementary, insofar as humans might be able to estimate distances by processing both static and dynamic information, in order to respond to the needs of different perceptual situations in which only a subset of cues might be available.

In recent years, several studies have investigated the capacity to process both static and dynamic information in order to estimate distances. However, these studies have often relied on the strong hypothesis that humans are able

* Corresponding author. Address: IHPST, CNRS/Université Paris 1 Panthéon-Sorbonne, 13 Rue du Four, 75006 Paris, France. Fax: +33 144278647.

E-mail address: Matteo.Mossio@ens.fr (M. Mossio).

¹ These authors contributed equally to this research.

not only to estimate distances by processing information from different sources, but also to build an “abstract” representation of distances supposed to be independent from the perceptual context in which it was built, and then operational in new contexts involving different perceptual cues (Frenz & Lappe, 2005; Sun et al., 2004).

In Section 1, we review and discuss the main findings in the literature related to the following issues: Are humans able to estimate traveled distances accurately? Are traveled distances encoded as abstract representations? What is the specific role of optic flow in the estimation of traveled distances?

1.2. An abstract encoding of distances?

In recent years, extensive investigations have been conducted to test humans’ capacity to estimate distances by processing different categories of cues (visual or non-visual, static or dynamic). Within the same perceptual context, i.e. when the same cues are available during both the presentation (encoding) and the response (test) phase, the estimation of traveled distances can be very accurate. For instance, protocols using only dynamic non-visual information (idiothetic cues)—vestibular and somatosensory information combined or not with proprioception and efferent copies of an active locomotion—have clearly shown that idiothetic information is sufficient to accurately reproduce traveled distances (Berthoz, Israel, Georges-Francois, Grasso, & Tsuzuku, 1995; Bigel & Ellard 2000; Glasauer, Amorim, Vitte, & Berthoz, 1994; Harris, Jenkin, & Zikovitz, 2000; Israel, Grasso, Georges-Francois, Tsuzuku, & Berthoz, 1997; Mittelstaedt & Mittelstaedt, 2001; Sun et al., 2004). Similarly, very accurate estimation was observed with protocols in which static visual information was available in both the encoding and the test phase (Sun et al., 2004). These studies considered performance to be “accurate” when estimation errors were close to zero.

As Sun et al. point out, however, the interpretation of these findings is not straightforward: it is well known that the observed accuracy may hide intrinsic distortions of distance perception (due to the processing of a specific set of cues), which can be canceled out overall because both the encoding and the test phase occur in the same perceptual context. Hence, these protocols cannot provide unambiguous information about the intrinsic accuracy of the distance perception mechanisms. However, in most cases the estimation of traveled distances *does* occur in similar perceptual contexts, in which the effects of hypothetical distortions would not appear. A common way to address this issue has been to design cross-modal protocols in which cues belonging to different perceptual modalities are available in the encoding and test phases. The most striking finding of these studies is the increased estimation errors in the observed performance (see Table 1 for a review of recent results), together with a strong inter-individual variability.

The common interpretation of these results is twofold. First, humans possess efficient mechanisms to build an abstract representation of distances, which can be trans-

ferred from one perceptual context to another. Second, inaccuracy is mainly due to perceptual distortions produced by the specific set of available cues. For instance, as Sun et al. suggest, the inaccuracy observed in cross-modal protocols involving optic flow could be explained by the intrinsic underestimation of traveled distances due to the processing of dynamic visual information. Such an underestimation would lead to an undershoot of distances if optic flow were available during the encoding phase and an overshoot if it were available during the test phase.

An empirical argument supporting this interpretation comes from the observation that symmetrical cross-modal protocols (in which the perceptual context is reversed in the encoding and the test phases) lead to symmetrical results, i.e. the size of error is constant but inverted in sign (see Table 1, protocols 3 and 4). These results seem to suggest that humans are indeed able to build an abstract representation of traveled distances, since the specific sequence of the perceptual contexts does not affect the error size. At the same time, exposure to a given set of cues generates a characteristic misperception of distances, which has predictable empirical consequences on performance.

Nevertheless, some difficulties arise with this interpretation. First, the predicted symmetry in the results is not observed for all symmetrical protocols (see Table 1, protocols 1 and 2). Second, concerning the specific case of optic flow, protocol 5 results are inconsistent with the prediction that exposure to optic flow would lead to an underestimation of perceived distances, since reported results show the opposite tendency. Therefore, we hold that characteristic perceptual distortions produced by specific sets of cues cannot provide a completely satisfactory explanation for observed inaccuracy.

In this paper, we suggest a complementary explanation based on two working hypotheses. First, we propose that estimation errors could stem from a fundamental inefficiency in constructing an abstract representation of distances, independent from the specific perceptual transformation, given the little adaptive need for such a capacity. In the specific case of optic flow, since vision is most of the time available during self-motion, rare are the situations in which a context-independent representation is required. Accordingly, we suggest that the errors observed in cross-modal protocols involving optic flow are mainly the consequence of inefficient mechanisms dealing with unusual and non-ecological perceptual situations. This approach could provide an intuitive explanation of why protocols involving optic flow in the same experimental phase (see Table 1, protocols 3, 5, 7 and 8) obtain different performances. Rather than being a consequence of the optic flow processing distortions, these differences would result from different perceptual *transformations* between the encoding and the test phase. We do not claim that these intrinsic distortions do not exist, but suggest rather that a more elaborate explanation could account for the large errors observed in these protocols (for further details, see Section 5). Second, previous studies do not pro-

Table 1
Overview of the results reported in recent studies using cross-modal protocols

	Perceptual cues		Ratio testing/learning	References
	Learning phase	Testing phase		
1	Static visual	Idiothetic	Undershoot ($\cong 0.7$)	Sun et al. (2004) ^a
2	Idiothetic	Static visual	$\cong 1$	Sun et al. (2004)
3	Static visual	Dynamic visual + idiothetic	Overshoot ($\cong 1.3$)	Sun et al. (2004)
4	Dynamic visual + idiothetic	Static visual	Undershoot ($\cong 0.7$)	Sun et al. (2004)
5	Vestibular	Dynamic visual	Undershoot (0.23)	Harris et al. (2000)
6	Static visual (virtual cues)	Vestibular	Overshoot ($\cong 2$)	Harris et al. (2000)
7	Static visual	Dynamic visual + vestibular	Overshoot (3.57)	Harris et al. (2000)
8	Vestibular	Dynamic visual + vestibular	Overshoot (2.33)	Harris et al. (2000)
9	Static visual	Proprioceptive	Overshoot ($d < 15$ m)	Harris et al. (2002)
			Undershoot ($d > 15$ m)	

^a Blind Walking Task Paradigm, see also Mittelstaedt & Mittelstaedt (2001), Bigel and Ellard (2000), Fukusima, Loomis, and Da Silva (1997), Loomis, Da Silva, Fujita, and Fukusima (1992), Rieser, Ashmead, Talor, and Youngquist (1990), Steenhuis and Goodale (1988), Elliott (1987), Corlett, Patla, and Williams (1985), Thomson (1983). It should be noted that the results vary considerably following the specific experimental conditions. In particular, walking speed plays a crucial role in determining the observed performance.

vide any explanation for the usually large inter-individual variability, which suggests that humans do not rely on the same mechanisms in order to execute the task. We therefore investigated whether the great inter-individual variability reported may have stemmed from the inhomogeneity of the population from which subjects were selected. In the light of these hypotheses, we focused on the role of optic flow in the perception of traveled distances, by investigating subjects' reaction to different manipulations of visual cues, when optic flow was present in both the encoding and test phases.

In the next section, we review the recent findings on the contribution of visual cues to the estimation of traveled distances.

1.3. The role of visual cues

With the emergence of virtual reality (VR) and greater control over visual stimulations, several studies have focused on the contribution of visual cues (both static and dynamic) in the perception of traveled distances, by dissociating them from idiothetic cues. Different combinations of visual cues in the encoding and test phases have already been studied: static cues in the presentation followed by a dynamic reproduction; a dynamic presentation followed by a test with static cues; and a dynamic presentation followed by a dynamic reproduction.

When participants were asked to reproduce a previously presented static distance, with an optic flow translation simulated at constant speed or with a low acceleration, they significantly overshoot large distances (40% on average for distances of 4–32 m, Frenz & Lappe, 2005; Redlick, Jenkin, & Harris, 2001). In contrast, for shorter distances (under 4 m) subjects tended to undershoot by about 36%. Symmetrically, when they were asked to indicate statically distances they learned dynamically, subjects tended to undershoot distances by about 27% (Frenz & Lappe, 2005).

As with cross-modal protocols, these results could suggest that exposure to optic flow generates an underestima-

tion of traveled distances. However, this interpretation cannot fully account for the experimental results. A significant correlation between stimulus duration and the size of the estimation error was also reported: this could suggest that travel duration significantly modulates the underestimation, which would produce an overshoot only for longer distances. Furthermore, a strong inter-individual variability was observed. Here again, we hold that inefficient mechanisms to generate context-independent representations of distances could explain these observations better than a simple intrinsic misperception of distances using optic flow. Similarly, the inter-individual variability could be a consequence of substantial individual differences in the capacity to accomplish the task.

In order to test whether performance could improve when both the encoding and test phases involve dynamic visual information, requiring no abstract representations, other studies, based on a discrimination or a reproduction task, have used optic flow-simulated translations. In these conditions, the distance underestimation effects with optic flow should not be observed. Yet, according to the task, experimental results are still very different.

On the one hand, when subjects were asked to discriminate between two traveled distances, in which the velocity profile and duration varied, they were very precise, with an error rate inferior to 3% (Bremmer & Lappe, 1999; Frenz, Bremmer, & Lappe, 2003). As the authors emphasized, even if the performance was partially influenced by the visual environment in which the reproduction occurred (manipulation of texture and depth cues), subjects could discriminate distances very accurately, compensating the changes in the velocity profile between encoding and test. Once again, however, reported inter-individual differences in performance, despite their reduced error size as compared to cross-modal protocols, call for a finer account of experimental results.

On the other hand, when the task is to reproduce traveled distances, subjects significantly overshoot distances of up to 8 m and undershoot larger distances (Bremmer &

Lappe, 1999; Frenz & Lappe, 2005). Furthermore, performance is increasingly impaired with the complexity of the velocity profile experienced during the encoding phase (Bremmer & Lappe, 1999). Since the overestimation resulting from the processing of optic flow should not be observable in protocols involving the same perceptual context in both the encoding and test phases, these results are in sharp contradiction with the notion that the intrinsic misperception of optic flow is responsible for the poor performance reported.

In turn, we believe that the difference between discrimination and reproduction protocols could also be explained by the active nature of the reproduction task, which does not exist in the entirely passive discrimination protocols. In reproduction protocols, subjects had not only to update their estimate of the traveled distance, but also to control the instantaneous speed of the translation, which might have impaired performance. Moreover, each reproduction in the same condition could involve different velocity profiles, which could have been responsible for the not-so-small variability observed. Again, we maintain that inefficient mechanisms for perceptual transformations—in this case from a sensory to a sensorimotor mode—could explain the poor performance observed, and the variability could stem from the large individual differences in the strategy used to perform the task.

1.4. Objectives of the present study

In this study, we directly addressed the problem of understanding if, and under what conditions, humans are able to process information provided by optic flow to estimate traveled distance accurately. Our general goal was to validate the working hypotheses we formulated in Section 1.2, according to which (1) the construction of a context-independent measure of traveled distance should be seen as an effort to adapt to particular conditions for which we do *not* have the adequate capacity and (2) the previously reported inter-individual variability when estimating traveled distances may result from substantial individual differences. Accordingly, we sought to show that performance could be significantly more accurate and homogeneous if (1) the estimation took in similar perceptual contexts involving optic flow and (2) subjects were grouped according to their performance and following a certain criterion.

To achieve this goal, we designed an experimental protocol in order to assess the following three following fundamental and closely interconnected questions:

- (1) What degree of *accuracy* and inter-individual *homogeneity* can be attained by humans when they can process information provided by optic flow (which is most of the time available in real displacements) if no abstract representations are needed?
- (2) How do humans actually process the information, i.e. what level of *cognitive resources* is required in order to correctly estimate traveled distances?

- (3) What exactly is the relevant information needed to accurately reproduce traveled distances, given that optic flow contains several cues that could potentially be processed?

In fact, the general hypothesis according to which optic flow can be processed to estimate traveled distances needs further specification, since many sources of information could potentially be used by humans to accomplish the task, and many perceptual mechanisms could be involved. Several perceptual strategies can be adopted to estimate traveled distances but, as we shall see, none of them was unambiguously supported by the experimental data.

2. Materials and methods

2.1. Participants

Twenty-two subjects, mostly students, participated in this study (12 males and 10 females). Their ages ranged from 22 to 37 years. All had normal or corrected-to-normal vision.

2.2. Apparatus

Participants sat on a chair wearing a helmet-mounted display (SEOS HMD 120/40) in a completely dark room (see Fig. 1). Through the HMD, they experienced a three-dimensional visual environment subtending $120^\circ \times 64.4^\circ$ of horizontal \times vertical field of view, with a 40 overlap for

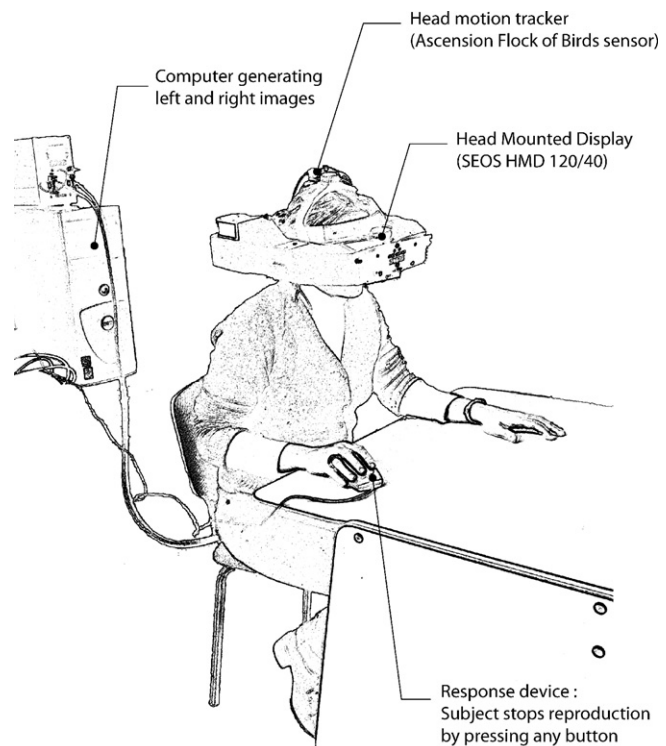


Fig. 1. Illustration of the experimental set-up. The virtual translations were projected through a stereoscopic helmet-mounted display subtending 120° of horizontal field of view with 40° of overlap, and a resolution of 1280×1024 for each eye. A head-tracker was used to reproduce head movements in the rendered environment and responses were given with a mouse.



Fig. 2. Static view of the virtual environment as experienced by the subjects.

stereoscopic vision. Subjects were free to perform all physical head translations and rotations. A motion tracker (Ascension ‘Flock of Birds’) recorded these head translations and rotations, which were then processed in order to reproduce in the simulated environment the corresponding visual motion. The virtual environment was a linear portion of a street. The initial viewing position in this street was always placed at a height of 1.70 m above the ground, and equidistant from the two sidewalks to the right and left side (see Fig. 2). The viewing position height was kept constant during the whole experiment. A computer generated the real-time simulation of passive forward translations along the street. All participants reported having clearly understood and perceived that they were being displaced in this virtual street.

2.2.1. Manipulation of the cues available for the reproduction

The experimental setup allowed us to manipulate two general sets of perceptual cues potentially involved in the estimation of traveled distances. Cues were manipulated in the test phases only, therefore the encoding phases were all identical except for the traveled distances.

Firstly, we manipulated travel cues, i.e. those cues that are produced by the optic flow. In particular, we distinguished between temporal and spatial travel cues (see Fig. 3, upper panel). The *Temporal travel cues* we manipulated were velocity and duration: traveled distances can be estimated in terms of duration if travel velocity is invariable between the encoding and the test phase. In order to test their contribution, we manipulated the velocity profile, by replacing the constant speed (1.2 m/s) used in the encoding phase with three different constant accelerations (0.5, 0.7 and 0.9 m/s²) during the test phase. The constant velocity was chosen in order to simulate a natural walking speed. The accelerations were chosen so to avoid strong accelerations and to prevent having equivalent travel durations for the tested distances for constant and accelerated velocity profiles. *Spatial travel cues* could also be potentially involved in the estimation of traveled distances: subjects might thus estimate traveled distances by exploiting texture regularity between the encoding and test phases, i.e. by quantifying textural units traveled. In our protocol, each element of the environment (roadway, sidewalks, walls) had a specific regular texture, which differed from each other with respect to their spatial frequency (see Fig. 2). When texture was manipulated in the test phase, the regular texture units were eliminated and replaced by dot-like units, so to obtain a “sand effect”, thus making it impossible for participants to estimate distances by counting the scrolling spatial units. This manipulation suppressed the low-frequency content of the scene provided by the different regularities, but also decreased slightly the overall contrast. In order to avoid flickering textures, the mip-mapping texturing technique was used. Consequently, the textures, whether regular or dotted, were blurred in depth beyond approximately 4 m of distance. This prevented subjects from using a strategy relying on spotting precise landmarks at the estimated traveled distance, without altering the forward motion information provided by optic flow.

A second general category of cues labeled “depth cues” was manipulated (see Fig. 3, middle panel). These cues were manipulated *jointly* in some of the experimental conditions so to investigate their global contri-

bution to the estimation of traveled distances. The depth cues available in our protocol were *lateral parallax*, generated by lateral head movements, and *stereovision*, produced by the binocular disparity between the images rendered for each eye with the stereoscopic HMD. A third depth cue, *perspective*, was provided by the global layout of the environment and its intrinsic geometric structure. During the test phase, lateral parallax was manipulated by suppressing the transmission of lateral head movements in the calculation of the camera position in the virtual environment. The rotational head movements were always transmitted to the camera orientation in the virtual environment. In this condition, during the simulated translations, the physical head translations of subjects recorded by the head-tracker were not reproduced in the simulated environment. Stereovision was manipulated by suppressing the binocular disparity between the two cameras, and displaying the same viewpoint in both eyes. Perspective was manipulated by modifying the geometric structure of the environment: the walls, which were parallel in the encoding phase, were bent so to converge in the test phase. We chose an arbitrary distortion corresponding to a street width reduction of 50% after a 100 m, which was clearly visible and would induce large differences for distance estimation.

These depth cues are usually involved in the static perception of distances, and they are independent from the forward translation of the observer through the environment. It should be noted that, whereas lateral parallax and stereovision provide absolute information about distances, travel cues are intrinsically relative. However, since in our protocol the viewing position height and the size of the texture units were kept constant, no scaling problem has to be solved to estimate distances.

2.2.2. Dual task

In order to test and limit the contribution of high-level cognitive resources to task performance, we introduced a dual task in the test phase of all the conditions but one, which served as a reference condition. A number between 40 and 60 was displayed before the test phase. Participants were instructed to start subtracting 4 from the displayed number, and to say each new result aloud at approximately 2-s intervals until they ended the reproduction (see Fig. 3, lower panel). The experimenter was constantly checking if the dual task was being executed carefully: participants made very few mistakes in counting, and we did not have to eliminate any trials in which there was a counting error. We chose to introduce this dual task in order to specifically test to what extent explicit verbal counting strategies of spatial or temporal travel cues are required in order to accurately estimate distances.

2.3. Procedure

2.3.1. Time-course of a trial

The experimental procedure consisted in a sequence of trials composed of an encoding phase followed by a test phase. During the encoding phase, participants saw an initially static view of the virtual environment and, after a delay of 2 s, the simulated self-motion started automatically. Participants were then passively translated at a constant speed (1.2 m/s) over the tested distance along the street, and then the simulated motion stopped

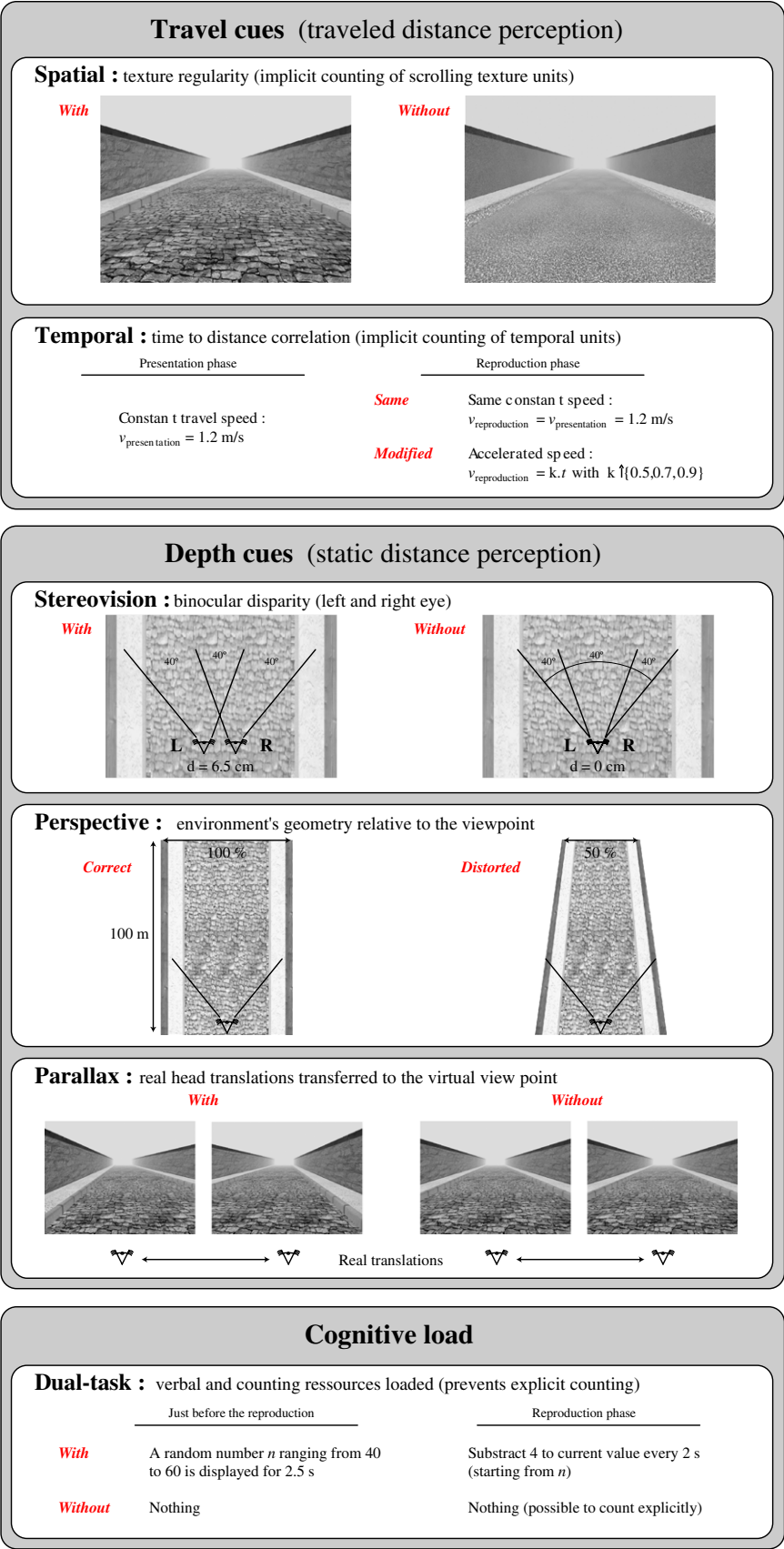


Fig. 3. Illustration of the studied factor manipulations according to each category: depth cues (stereovision, perspective and parallax), travel cues (spatial and temporal), and cognitive load (specifications of the verbal dual task).

Table 2
Experimental conditions with their corresponding factor manipulations

	Condition	Cognitive load	Manipulated factors		Temporal travel cues
			Depth cues	Spatial travel cues	
1	<i>Control without DT</i>				
2	<i>Control with DT</i>	×			
3	<i>Depth cues</i>	×	×		
4	<i>Texture</i>	×		×	
5	<i>Velocity</i>	×			×
6	<i>Velocity & texture</i>	×		×	×
7	<i>All cues</i>	×	×	×	×

2 s before the environment disappeared. All visual cues described above were always available during the encoding phase: accordingly, subjects were explicitly allowed to use whatever strategy they preferred in order to estimate the traveled distance. During the test phase, the virtual environment was again shown for 2 s before the simulated self-motion started. Participants were instructed to press any of the mouse buttons to stop the automatic forward translation as soon as they felt they had covered the same distance as during the encoding phase. In the test phase, depending on the specific experimental condition, only a subset of the visual cues was still available.

We chose to manipulate the cues in the test phase in order to assess their role in the construction of an internal representation of a traveled distance in ecological situations, where all cues are available. The starting position in the street was randomized in both phases so as to eliminate the use of landmarks taken from the textures.

2.3.2. Experimental design

We designed seven experimental conditions, which differed from each other according to the specific manipulation of cues. Each condition was repeated for three different distances (7, 9 and 12 m) and three times for each distance. Consequently, each participant repeated each condition nine times. The entire experiment thus included 63 trials (9 repetitions \times 7 conditions). The 63 trials were presented in random order. Except for the first condition we shall discuss, the dual task was always introduced in the test phase. Each participant was tested in every experimental condition. The total duration of the experiment was approximately 50 min. Participants were allowed to have a break between two trials every 10 min. After the end of the entire experiment, we asked the participants to describe the strategy they had employed to accomplish the task and how difficult it had been. The factor manipulations of each experimental condition are summarized in Table 2.

2.4. Data analysis

The only experimental performance that was collected and analyzed was the ratio between distances reproduced in the test phase and distances presented in the encoding phase, calculated for each trial. Three percent of trials were considered aberrant and were excluded from the statistical analyses: these were trials in which participants explicitly reported a loss of concentration (resulting, for example, in their confusing encoding and test phases), and outlier trials (where the result was more than two standard deviations from the mean of the condition in question).

Statistical analyses were done using a repeated-measures ANOVA design. Conditions were compared in pairs and, when necessary, a categorical group factor was introduced in the analyses ('velocity profile dependency', discussed later). Post hoc analyses were performed using the Tukey test. The only dependent variable was the reproduction ratio defined above.

3. Restricting the use of high-level cognitive strategies

In order to determine on which visual cues humans base their estimations of traveled distances, we need to ensure

that they do not use strategies relying on non-visual information. Therefore, we restricted the access to explicit verbal strategies by means of a dual task. In the experimental conditions discussed in this section, we will show how the dual task allowed us to identify and eliminate subjects possibly relying on such strategies.

3.1. Accuracy and cognitive load

Some people might possibly rely on high-level cognitive strategies in order to estimate traveled distance: in this case, reproducing distances would be much harder or even impossible without specific attentional resources and the use of verbal strategies. In contrast, others might rely on more automatic processes, which would operate independently of the availability of high-level resources. The second possibility seems to have some ecological plausibility: if humans possess the capacity to "automatically" estimate traveled distances, they might be able to do so without allocating much of their attentional and verbal resources. Unfortunately, existing protocols have not given any clear insight into this issue. Only one experimental protocol using only visual cues involved a dual task (Sun et al., 2004), and its effect on performance was not compared with a control condition performed without the dual task. Therefore, it is possible that high-level cognitive strategies played an important role in the performance reported, and the inter-individual variability could partly stem from very different individual strategies to perform the task. In order to test this assumption, we compared performance when the dual task had to be or did not have to be executed in parallel during the reproduction (conditions labeled *Control with DT* and *Control without DT*, respectively). In these conditions all the visual cues were available in both phases, and the speed was kept constant (1.2 m/s).

3.2. Results and discussion

In the *Control without DT* condition, the overall performance of all 23 subjects was very accurate, with an average undershoot for the reproduced distances of 6.9%. The inter-individual variance was rather low (standard deviation of 9.4%). We observed a range effect for longer distances (undershoot of 4.5% and 3.4% for 7 and 9 m, respectively,

compared to an undershoot of 12% for 12 m). These results suggest that in this condition humans can perform the task very efficiently. However, since in the *Control without DT* condition none of the available cues was manipulated during the test phase, subjects' performance was not very constrained. In particular, they could have used efficient strategies not relying on vision, such as explicit counting of time or textural units, in order to accomplish the task. Results in the *Control with DT* condition thus provided crucial additional information about the cognitive resources required. In the *Control with DT* condition, the overall performance of all 23 subjects was still very accurate even if we found an average overshoot of 6%. The inter-individual variance, in turn, increased substantially (standard deviation 19.6%, F -test: $F(44) = 4.34$; $p < .001$). In the *Control with DT* condition, some subjects could still perform the task with great accuracy (we again observed a range effect: undershoot of 1.1%, 5.2% and 12% for 7, 9 and 12 m, respectively) while others were severely disturbed by the introduction of the dual task in the test phase. We ran a cluster analysis to verify if two distinct groups of subjects could underlie these different behaviors. A tree diagram based on the Euclidian distances of subjects' performance (average reproduction ratio for each subject, pooled over distances) in both the *Control without DT* and *Control with DT* condition was computed using Ward's method (see diagram in Annex 1). The first level of branches in the tree-diagram distinguished between two groups of subjects. The first one included eight subjects whose mean reproduction ratio in the *Control with DT* condition increased by at least 12% with respect to the *Control without DT* condition, and who overshoot the traveled distance by at least 13% in the *Control with DT* condition. The second group was composed of the 15 remaining subjects, whose performance was still very accurate in the *Control with DT* condition, in terms of both mean ratio (undershoot of 5.91%) and standard deviation (7.61%). The cluster analysis therefore allowed us to dissociate subjects who were "disturbed" by the dual task from those who were "not disturbed".

The interaction between the condition (Control with and Control without dual task) and the group (disturbed or not) was significant ($F(1,21) = 32.925$; $p < .0001$). The Tukey post hoc tests confirmed that disturbed subjects had a significant increase of the average reproduction ratio ($p < .001$), reaching 128%, whereas performance was not significantly different ($p > .8$) for non-disturbed subjects, with 94% (see Fig. 4).

We interpret this distinction as resulting from different strategies. Subjects who were not disturbed by the dual task did not rely on strategies involving explicit verbal counting to accomplish the task. It is worth emphasizing that the results we report for these subjects are more accurate than those reported in previous cross-modal and visual studies. Accordingly, we state that the estimation of traveled distances can be very accurate if the perceptual context is the same in both encoding and test phase. Furthermore,

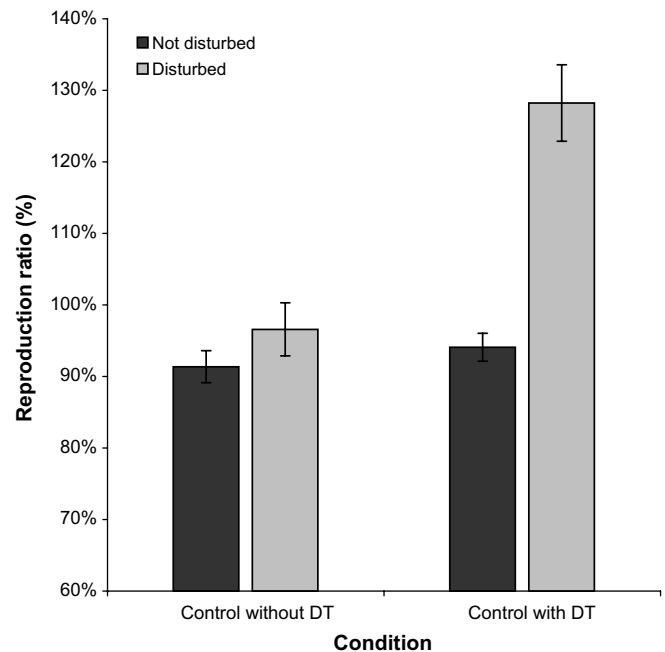


Fig. 4. Distance reproduction plotted as a function of the control condition (with or without dual task) and the group (disturbed or not). The error bars correspond to the inter-individual standard error.

for two-thirds of our subjects, this holds true despite the dual task.

The eight subjects whose performance was impaired reported at the end of the experimental session having been strongly disturbed by the dual task. Taken together, these observations suggest that these subjects were either relying on explicit verbal counting strategies in order to estimate traveled distances, or were unable to perform the main task together with the dual task (on this point, see Section 5). This is consistent with the performance of these subjects across all the tested conditions: they systematically showed a strong tendency to substantially overshoot distances, just as in the *Control with DT* condition. The influence of the dual task thus appears to prevail over the manipulation of all other perceptual cues. It is possible that these subjects, after a period of training in this quite unnatural situation, could eventually develop the capacity to estimate traveled distances despite the interference of the dual task. However, when we compared trials of the same experimental condition occurring at different moments of the protocol, we did not find any significant amelioration of the performance of those subjects. We conclude that, at least with respect to the protocol duration, no learning effect occurred.

Given the focus of our study on low-level and visual mechanisms, the results of these eight subjects in all the remaining conditions (where the dual task was always introduced) were excluded from further analyses. The *Control with DT* condition was used as the reference condition with which the others will be compared. In Section 5, we will argue that the inter-individual variability reported in previous studies could be—at least partially—explained

by a failure to distinguish between the two categories of subjects we have described in this section.

4. Which visual cues are processed to estimate traveled distances?

In the following sections, we address the question of how humans process optic flow in order to estimate traveled distances when the use of explicit verbal counting strategies is not allowed. Our general objective was to test the efficiency of several possible strategies that could be used in order to estimate traveled distances. All these strategies will be introduced and discussed successively.

4.1. Are depth cues necessary?

4.1.1. Depth summation strategy

The first perceptual strategy we tested, not yet explicitly explored in the literature, consists in estimating traveled distances through the same kind of mechanisms used to perceive depth. Accordingly, all along the path the observer could sum the length of successively traveled portions estimated statically using depth cues, which would ultimately provide an estimation of the global length of the traveled distance. If subjects applied this strategy, the estimation of traveled distances would not rely on travel cues, but rather on a summation of static estimations of distance. Indeed, once subjects have measured the distance of a point B from a specific location A, it would be sufficient for them to be placed at point B and then calculate the distance of a new point C from B. Although recent studies (Frenz et al., 2003) have argued that depth cues are not necessary to estimate traveled distance, several results have shown that performance is strongly influenced by manipulation of the available environmental cues, including depth cues. However, these results did not clearly distinguish the effect of depth cues from that of textural cues (see Section 5 on this point).

We designed the *Depth cues* condition to investigate whether subjects use the depth summation strategy. If depth cues were necessary to estimate traveled distances, their selective manipulation should significantly affect performance during the reproduction task. As detailed in Section 2.2, all depth cues were suppressed in the test phase; velocity was kept constant and textural cues were not manipulated. The dual task was maintained (see Table 2).

4.1.2. Results and discussion

For 15 subjects, the reproduction in the *Depth Cues* condition (undershoot of 6.3%) did not differ from that of the *Control* condition (non-significant main effect; see first and second bar of Fig. 5).

Performance in the *Depth cues* condition clearly showed that there was no influence of the manipulation of depth cues (stereoscopic vision, geometric perspective and parallax altogether) in the reproduction of traveled distances. Therefore, static cues, such as ocular and pictorial cues,

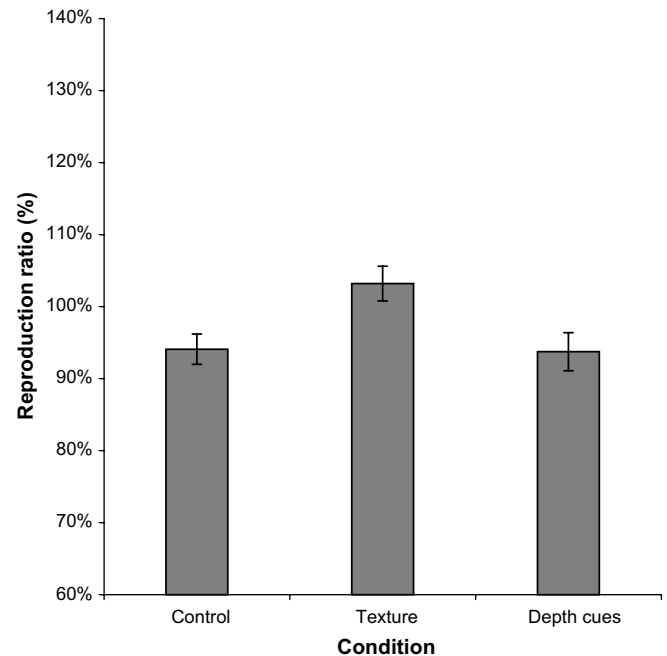


Fig. 5. Distance reproduction in the *Control*, *Depth cues* and *Texture* conditions (15 subjects). The error bars correspond to the inter-individual standard error.

are not necessary to estimate traveled distances, thus ruling out the possibility that subjects relied on the depth summation strategy to perform the task. A possible objection could be that the horizontal binocular overlap of our setup was limited to 40° as compared to approximately 120° for normal vision, which reduces the validity of our conclusion with regard to the stereoscopic vision. We argue that this objection is not compelling since, in the simplified environment we used, all relevant stereoscopic cues were available in the central zone of the visual field, and no additional information would have been provided by a larger overlap.

4.2. Is the invariance of spatial travel cues necessary?

4.2.1. Texture-unit counting strategy

A second possible strategy, already partially explored in the literature, consists in estimating distances by quantifying textural units, such as the cobblestones in the pavement, scrolling during the forward translation. If subjects adopted this strategy, we would expect a significant deterioration in the observed performance after the suppression, in the test phase, of the texture that had been available during the encoding phase. We designed the *Texture* condition to test whether subjects use the textural unit counting strategy. In this condition, we suppressed texture units as described in Section 2.2. In contrast, the velocity profile was kept constant and depth cues were not manipulated. The dual task was maintained (see Table 2).

4.2.2. Results and discussion

In the *Texture* condition, the mean reproduction ratio reached 103%, with a significant increase of 9.8% with

respect to the *Control* condition ($F(1, 14) = 9.97$; $p < .007$, see Fig. 5). The standard deviation increased slightly to 8.9%.

These results show that the manipulation of texture had a significant effect on the estimation of traveled distances, leading to a systematic overshoot (3%). The explanation for this effect could be that, in the absence of idiothetic information, the replacement of texture units with a uniform layout in the test phase reduced the contrast of the scene. It is well known that contrast reduction decreases the perceived motion speed (Thompson, 1982). Accordingly, subjects stopped the reproduction later than they did without the texture manipulation, which resulted in a significant overshoot. Although the suppression of texture regularity had a significant effect, this was rather weak compared to the reported effects produced by the other manipulations and did not prevent subjects from producing an accurate performance. This suggests that they were not relying primarily on the texture-unit counting strategy to perform this task. Therefore, we conclude that subjects' performance does not depend on the invariance of environmental texture units to estimate traveled distances, although it can be modulated by changes in the textural pattern.

4.3. Is the invariance of temporal travel cues necessary?

4.3.1. Strategy based on temporal travel cues

The *Texture* condition showed that, even if the suppression of texture regularity influenced the estimation of traveled distances, subjects' performance was still rather accurate. Hence, the outcome of previous *Depth Cues* and *Texture* conditions leaves open the question as to how subjects succeed in accurately estimating traveled distances. Bremmer and Lappe (1999) suggested that subjects rely on the invariance of the velocity profile between the presentation and the test phase to accurately reproduce traveled distances. Similarly, Frenz and Lappe (2005) have suggested that travel duration has a strong influence on the estimation of distances. In this view, traveled distances would be perceptually represented in temporal terms (elapsed duration) rather than in spatial terms (number of scrolled texture units).

In light of these interpretations, we designed the *Velocity* condition to test whether subjects' strategy could rely on temporal travel cues—in our case provided by the velocity profile—in order to estimate traveled distances. We replaced, in the test phase, the constant velocity by one of the three accelerated velocity profiles (as defined in Section 2.2). Each of the three accelerated profiles was tested for each condition and distance. If subjects' based their strategy on travel duration, manipulation of the *Velocity* condition would lead to predictable changes in performance. In this condition, for all distances, a shorter duration is needed to reproduce at accelerated speeds the same distances as those learned at a constant speed. Hence, if subjects reproduced traveled distances by reproducing

Table 3

Predicted performance for a strategy based on the reproduction of the travel duration, compared to the performance observed for VPD subjects

Distance (m)	Mean reproduction ratio over three accelerations	
	Duration-based prediction (%)	VPD subjects' performance (%)
7	170	133
9	218	120
12	292	109

the travel duration, a systematic overshoot should be observed. Table 3 summarizes these predictions ("Duration-based prediction" column). It is worth noting that the predicted overshoot should tend to increase with distance: the greater the distance, the larger the overshoot.

Anticipating that subjects might still perform accurately in the *Velocity* condition, we designed two further conditions (*Velocity & texture* and *All cues*) in order to determine whether subjects were in fact "switching" between different strategies and thus processing different cues, according to their availability in the tested perceptual context (see Table 2). The velocity profile was manipulated in the test phases for each of these three conditions.

4.3.2. Results and discussion

Performance of the 15 subjects in the *Velocity*, *Velocity & texture* and *All cues* conditions shared two common properties: a global overshoot of traveled distances and a substantial increase in inter-individual variance, which exceeded 20% in each condition. The increase in inter-individual variability led us to the hypothesis that the 15 subjects possibly relied on different strategies to estimate traveled distances. While looking at individual results, we observed two distinguishable behaviors: some subjects were still able to estimate traveled distances as accurately as in the control condition, whereas others considerably overshoot distances in all three conditions.

In order to verify whether two distinct groups of subjects could underlie these different behaviors, we again ran a cluster analysis based on the overall performance of the 15 subjects in all 6 conditions with the dual task (see the tree diagram in Annex 2). The first level of branches in the tree diagram distinguished between two groups, one with eight and the other with seven subjects, which corresponded, respectively, to the two distinct behaviors observed in the conditions with an accelerated profile. The same cluster analysis can project the distance between conditions according to the performance of the 15 subjects (see tree diagram in Annex 3). The first branch in the conditions' tree diagram distinguished between the *Control*, *Depth cues* and *Texture* conditions on the one hand, and the *Velocity*, *Velocity & texture* and *All cues* conditions on the other hand. The only factor that differed between the two groups of conditions was the manipulation of the velocity profile in the test phase, which was constant in the first group and accelerated in the second. This point

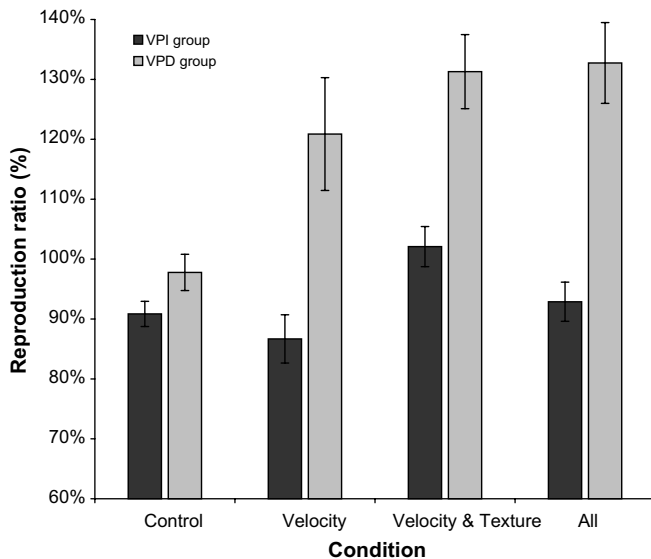


Fig. 6. Results for the accelerated reproduction conditions (*Velocity*, *Velocity & texture*, *All cues*) compared with the *Control* condition. All results are plotted for the VPI and VPD groups separately. The error bars correspond to the inter-individual standard error.

provides further evidence that what the cluster analysis has been discriminating is precisely the different reactions of subjects to manipulation of the velocity profile. Accordingly, we labeled the two groups of subjects “velocity profile independent” (VPI) and “velocity profile dependent” (VPD), respectively. The distinction between VPI and VPD subjects was striking in all three conditions in which the velocity profile was accelerated in the test phase (see Fig. 6). The group (VPI and VPD) \times condition interactions were all significant (*Control* vs. *Velocity*: $F(1,13) = 6.75$; $p < .025$; *Control* vs. *Velocity & texture*: $F(1,13) = 8.01$, $p < .015$; *Control* vs. *All cues*: $F(1,13) = 12.95$; $p < .005$). We will discuss separately the performance of VPI and VPD subjects in these three conditions.

VPI subjects systematically showed very accurate performance in all conditions with an accelerated velocity profile, as they did in the previous conditions. In the *Velocity* condition, the post hoc test did not reveal a significant difference with respect to the *Control* condition ($p = .9$), and the standard deviation remained rather small (11%). This indicates that VPI subjects were able to compensate very efficiently for changes in the velocity profile, and that the whole group showed very similar performance. Therefore, their strategy in this condition did not rely on the invariance of the velocity profile. We designed the *Velocity & texture* in order to test whether this good performance was due to a strategy adaptation to the sensory context, relying alternatively on the velocity profile in the *Texture* condition and on the texture regularity in the *Velocity* condition. In this condition, we manipulated jointly the texture regularity and the velocity profile, whereas depth cues were preserved. If VPI subjects were able to rely alternatively on spatial or temporal travel cues, performance should now be severely impaired. In the *Velocity & texture* condition,

though, their performance (102%, with a standard deviation of 9.5%) was not significantly different from the *Control* condition ($p = .2$). The reproduction ratio in the *Velocity & texture* condition was increased by 16% as compared to the *Velocity* condition, which seems to confirm that the manipulation of texture regularity modulates the performance, introducing a slight but noticeable overshoot. We concluded that VPI subjects did not rely on the invariance of travel cues (whether spatial or temporal) to accurately estimate traveled distances. In the *Velocity & texture* condition, VPI subjects could possibly have switched to another strategy, relying this time on the invariance of depth cues in order to cope with the unreliability of all travel cues. To test this rather unlikely rescue strategy, we designed the *All cues* condition, in which we manipulated jointly both travel cues and the depth cues. Again, VPI subjects’ performance was very accurate in terms of both reproduction ratio (92%) and standard deviation (9.2%). Performance was not significantly different from that of the *Control* condition. These results suggest that depth cues do not play a role, even when travel cues are manipulated. Overall, VPI subjects have shown the capacity to adapt to a wide range of manipulations affecting simultaneously travel and depth cues in order to produce very accurate estimations of traveled distances.

In contrast, VPD subjects showed a completely different behavior in all conditions where the velocity profile was accelerated. In the *Velocity* condition, they substantially overshoot traveled distances by 20%, and the standard deviation doubled (24.9%). The post hoc test confirmed a significant difference between the *Velocity* and the *Control* conditions ($p < .05$). This overshoot is broadly consistent with the hypothesis that the subjects’ estimation was based on the perception of the travel duration, which indicates that VPD subjects relied on the invariance of the velocity profile. However, as Table 3 summarizes, the observed overshoot (33%, 20% and 9% for 7, 9 and 12 m, respectively) was largely inferior to that predicted with the acceleration profiles (70%, 118% and 192%). Since the speed started immediately at 1.2 m/s in conditions with a constant profile and started from zero with accelerated profiles, subjects could immediately detect that the velocity profile was accelerated. Hence, they possibly tried to compensate for this acceleration, stopping the reproduction before what the pure time-based strategy would predict. Accordingly, the greater the distance, the faster the traveling speed would become, resulting in an urge to stop the reproduction earlier. This interpretation would account for the smaller overshoot observed for greater distances, for which the compensating strategy works better. Nevertheless, the compensation was not fully efficient and not identically tuned across subjects, as shown by the increase in inter-individual variability. In the *Velocity & texture* and *All cues* conditions, VPD subjects’ performance confirmed the findings resulting from the *Velocity* condition. Distances were also significantly overshoot (by 31% and 32.7%, respectively) as compared to the *Control* condition,

with a large variability across subjects (standard deviation of 16.4% and 17.8%, respectively). In conclusion, VPD subjects' strategy cannot adequately cope with the manipulation of the velocity profile, since their performance in this condition is clearly impaired as compared to that of VPI subjects.

5. General discussion

This study investigated the contribution of optic flow to the estimation of traveled distances. Our main findings concern four issues: the kind of cognitive resources needed to estimate traveled distances; the great accuracy in the estimation shown by most subjects; the minor contribution of depth cues and texture regularity; the reliance of some participants on velocity profile invariance in contrast to the observed capacity of others to reproduce traveled distances even when the velocity profile is manipulated.

5.1. The dual task affected one participant out of three

The first contribution of the study concerns the role of high-level cognitive capacities in the estimation of traveled distances. When we manipulated the cognitive load in the two control conditions, we found that the performance of eight subjects (35% of the sample) was severely impaired by the introduction of the verbal dual task. For those subjects, the accurate estimation of traveled distances seems to require the use of high-level cognitive resources, such as the explicit counting of spatial or temporal travel cues. As already pointed out, we cannot rule out the possibility that some of these subjects, instead of relying on explicit counting, were in fact unable to handle efficiently the interaction between the main and the dual task. In order to exclude all the members of the first category of subjects, we had to accept the eventuality of excluding some members of the second one. For any of these disturbed subjects, the dual task possibly produced an additional cognitive load which they could not handle, and that interfered with the main task resulting in the substantial overshoot of traveled distances observed (see Section 3.2). It was recently proposed that an arithmetical dual task influences the perception of durations and therefore the distance reproductions (Glasauer, Schneider, Grasso, & Ivanenko, 2007). In particular, Glasauer and coworkers showed that distance reproductions decrease or increase by about 25% if the dual task is introduced in the encoding or test phase, respectively. This finding is to some extent consistent with the performance of our eight "disturbed" subjects. Nevertheless, the remaining subjects showed no noticeable increase of reproduced distances when introducing the dual task, suggesting that its interference becomes significant only for subjects who need to allocate important high-level cognitive resources to execute the dual task.

Moreover, this finding could provide a clue to explain the variance often reported in previous studies, which did not constrain the use of cognitive resources to perform

the task and were unable to distinguish between the different cognitive strategies participants may have relied on. Indeed, the variety of cognitive strategies could have led to differences in performance, thereby introducing great inter-individual variability. In the study by Sun et al. (2004), who also used a distractor task in their paradigm, participants whose performance was affected by the dual task were not considered as a separate group, which could have been a possible source of inter-individual variability.

An interesting finding was the absence of motion sickness when the protocol included the dual task in the test phases. Indeed, in a pilot experiment without the dual task, all participants reported motion sickness, some of them to the point that we had to interrupt the experiment before the end. In contrast, most of the participants in our study did not feel motion sickness with the dual task, and all of them were able to complete the experiment. The commonly accepted explanation for motion sickness is based on sensory conflicts (Reason, 1978; Reason & Brand, 1975). Sickness would be related to the incoherence between the information about self-motion provided by different sensory sources. In our study, it seems that the influence of the sensory conflict inherent in virtual reality, from which motion sickness might originate, was strongly reduced when cognitive and attentional resources were constrained. In the pilot experiment, where no dual task was used, many subjects tended to move their head freely and direct their gaze toward different components of the environment, so to estimate distances by explicitly counting texture units. In contrast, after the introduction of the dual task, most subjects stopped using this strategy, as shown by their reduced head movements, and tended to stabilize the gaze direction toward the focus of expansion. This stabilization could have markedly reduced motion sickness, because they never dissociated the gaze direction and the simulated heading direction, as they were free to do throughout the experiment.

5.2. General accuracy of performance

The second contribution of the paper concerns subjects' accuracy in estimating traveled distances. We found that 15 participants (65% of our sample) were able to produce a very accurate performance even after the introduction of the dual task. For them, the estimation of traveled distances did not require the use of explicit verbal counting strategies.

These findings contrast with the much greater inaccuracy reported in previous reproduction protocols involving optic flow in both the encoding and the test phase (Bremmer & Lappe, 1999; Frenz & Lappe, 2005). The existence of intrinsic distortions generated by the exposure to optic flow does not adequately explain the inaccuracy observed with reproduction protocols. Indeed, since optic flow was present in these protocols in both the encoding and the test phase, and no cue was manipulated, the visual distortions should be canceled out. The main difference with our protocol is the active control of motion speed. We used a

passive task where subjects were only asked to press a button in order to stop the forward motion. In contrast, in reproduction protocols, subjects had not only to estimate traveled distance but also to take into account the instantaneous motion speed, a situation that has two fundamental consequences. First, active reproduction tasks involve sensorimotor loops, which could interfere with the perceptual mechanism involved in performing the task. Second, the velocity profile in the test phase changed from one subject to another, which could have been responsible for the higher inter-individual variability reported in these studies. This interpretation is also consistent with the overall low level of errors reported in the discrimination protocols in which the test phase and the encoding phase were identical (Bremmer & Lappe, 1999; Frenz & Lappe, 2005). However, their strong inter-individual variability calls for a finer explanation, a point that we shall discuss in Section 5.4.

Of course, we do not exclude possible perceptual distortions that are intrinsic to the processing of optic flow. As mentioned in the introduction, these distortions could account for previous experimental findings, such as the observed symmetrical performance obtained with symmetrical protocols (see Table 1, experiments 3 and 4). Moreover, a “leaky integrator model” was recently proposed to account (at least partially) for results obtained in protocols alternating static and dynamic visual stimulation in the encoding and test phases (Lappe, Jenkin, & Harris, 2007). According to this model, optic flow integration contains an intrinsic leak factor whose effect on distance estimation changes depending on the specific experimental condition. In this case, the misperception generated by the leakage would not be canceled out in conditions involving optic flow in both the encoding and testing phase, and could then partially account for the inaccuracy observed in these protocols. However, it is not clear how this model would explain very accurate performance, as we report here. Moreover, the model predicts accuracy decreasing with the tested distances, which cannot account for the fact that the performance we report for distances from 7 to 12 m is as accurate as that reported by Frenz and coworkers for distances from 1 to 6 m. The different results seem to stem mainly from the specific transformation required. In fact, we believe that an intrinsic misperception (possibly due to a leak factor) could interact with the inefficient mechanisms in constructing context-independent representations. Since each specific perceptual transformation would rely on distinct mechanisms, the effects of intrinsic misperceptions could be enhanced, reduced or even reversed, according to the efficiency of the specific mechanism. Future experimental research should seek to describe the mechanisms underlying each specific transformation, and clarify how they interact with the intrinsic misperception of the involved modality. This explanation could account for the results obtained with protocols involving static and/or dynamic visual cues, but also for the cross-modal protocols mentioned in Section 1.

5.3. Depth cues and spatial regularity are not necessary

The third general finding of our study is that the accuracy of performance is largely unaffected by the manipulation of depth cues and texture regularity. Previous studies did not selectively manipulate the depth cues we considered (stereovision, lateral parallax and perspective) to assess their role in the estimation of traveled distances. In our protocol, we found that humans did not rely on depth cues to estimate traveled distances, regardless of whether travel cues were available or not (invariance of velocity profile and texture regularity). This suggests that the perception of traveled distances and the perception of depth rely on distinct mechanisms.

Conclusions on the contribution of texture regularity are less straightforward. We found a significant effect of the manipulation of texture regularity on the estimation of traveled distances. The observed overshoot of traveled distances depended on the specific texture changes we chose: reducing the contrast of the scene possibly led to an underestimation of travel speed (Thompson, 1982). The suppression of low-frequency spatial regularities resulting from the same manipulation could possibly have an influence in the performance. Indeed, the influence of spatial frequency on motion perception has long been known (Pavard & Berthoz, 1977). However, since the accuracy of reproduction was still remarkable, we conclude that the estimation of traveled distances does not rely on the implicit counting of scrolling textural units, even though texture regularity slightly improves performance. This is consistent with the findings of previous studies. Bremmer and Lappe (1999) reported that texture manipulations had significant effects on the reproduction and discrimination of distances. However, since textural cues were manipulated simultaneously in the encoding and test phases, they tested the capacity of the perceptual mechanisms to cope with impoverished visual environments, rather than the specific role of natural textures available in ecological environments. More recently, Frenz et al. (2003) found no significant effect when replacing the textured ground plane of the encoding phase with a dot plane in the test phase. The difference between their findings and ours could stem from distinct manipulations of textures. Despite the minor differences reported by these studies, the invariance of texture regularity was not required for accurate performance.

5.4. Importance of the velocity profile

Our fourth and main result concerns the role of velocity profile invariance in the estimation of traveled distances. We evidenced two distinct behaviors in response to manipulation of the velocity profile. Some participants were able to accurately estimate traveled distances regardless of changes in the velocity profile—and thus of travel duration—between the encoding and the test phase. In contrast, other participants appeared to be dependent on the invariance of the velocity profiles since their performance was

severely impaired by this manipulation. We suggest that the latter made use of an alternative strategy to try to cope with the accelerated profiles, which was not efficient in compensating for the changes in travel duration. This interpretation is supported by the increased inter-individual variability in this group, indicating differences in the efficiency of these compensatory responses.

These findings shed new light on the results reported in previous reproduction protocols. As previously reported for a purely vestibular reproduction of distances (Israel et al., 1997), Bremmer and Lappe (1999) claimed that humans reproduce the visual velocity profile when they are asked to actively reproduce distances. Nevertheless, an unexplained strong variability was observed. Only 63% of subjects actually reproduced the velocity profile, whereas 15% reproduced at a constant speed and 22% reproduced using other velocity profiles. Since subjects were free to choose any strategy to reproduce distances, this protocol could not probe whether subjects who reproduced distances using the same velocity profile could still do so with another velocity profile. In this sense, the results were not only variable, but they also crucially failed to clarify whether invariance of the velocity profile was required for accurate reproduction of traveled distances, at least for some subjects. Our experimental design allowed us to assess this issue. Indeed, we found that about one third of our participants relied on the invariance of the velocity profile to reproduce traveled distances. In future research it would be interesting to try to establish a direct link between this group of participants and those who, once they control the travel speed in the test phase, would reproduce the same velocity profile as in the encoding phase.

Furthermore, the variable reaction to manipulation of the velocity profile enables us to refine the interpretation of results obtained in discrimination protocols, in which subjects were asked to discriminate between traveled distances when velocity and duration varied simultaneously (Frenz et al., 2003). While our results are consistent in terms of accuracy with those measured with discrimination protocols, our analysis provides an account for the strong inter-individual variability reported (see Section 1.3). We believe that these protocols failed to distinguish between velocity profile dependent and velocity profile independent subjects. Indeed, velocity profile independent subjects would be able to discriminate between distances even if the speed and duration were changed, whereas this would be extremely difficult for velocity profile dependent subjects, who would be unable to compensate for the simultaneous variation of speed and duration.

The most surprising finding was the capacity of some subjects to produce an extremely accurate performance in conditions where visual cues were manipulated separately, or even simultaneously, while they were executing the verbal dual task. This capacity would require very efficient adaptive mechanisms that enable accurate estimations when dynamic visual information is available, independent of the simultaneous manipulation of the velocity profile,

texture regularity and depth cues. Therefore, we conclude that these subjects are able to estimate traveled distances by directly integrating the optic flow generated by the forward motion.

Is there independent evidence that visual systems could directly process various image velocities through various transformations of the visual patterns? Indeed, experimental studies on the extrastriate middle temporal visual area (MT) in the macaque have shown that MT neurons are able to encode the direction of moving objects (Albright, 1984) and respond selectively to both the temporal and spatial frequency of the visual stimulus (Foster, Gaska, Nagler, & Pollen, 1985). More recently, several studies have investigated the hypothesis that the MT neural network encodes image speeds with a population of neurons responding to specific speeds, regardless of the spatiotemporal frequency of the visual stimulus (Perrone 2005; Perrone & Thiele, 2001; Priebe, Cassanella, & Lisberger, 2003). Experimental results seem to suggest that some MT neurons do indeed possess the capacity to respond to specific optical speeds with some degree of independence from the variation of spatiotemporal frequency. The existence of neurons able to respond selectively to specific optical speeds, independently of the spatiotemporal frequency, could provide clues to understanding the mechanisms underlying the good performance of the velocity profile independent subjects in our study.

6. Conclusions

In the present study, we simulated a virtual environment to test the contribution of optic flow to the estimation of traveled distances. We chose to use a virtual reality setup to control and restrict the available sources of information, in order to manipulate independently the visual cues, and to disentangle their respective contribution to the estimation of traveled distances. The virtual environment was designed to maximize its ecological plausibility, providing participants with a large field of view, stereoscopic vision, lateral parallax, free head movements, and realistic texturing and geometric structure. We showed that humans have the capacity to estimate distances in perceptual contexts where only optic flow is available, independently from the invariance of textural cues, depth cues and, for some subjects, even from the invariance of speed and duration. The experimental manipulations used to obtain these findings would not have been possible with physical setups, clearly illustrating the new possibilities offered by virtual reality to investigate spatial cognition.

Acknowledgments

The authors thank Michel Ehrette, Pierre Leboucher and Yves Dupraz for technical assistance and all the participants in the experiment. Matteo Mossio received a Ph.D. fellowship from the French Ministry of Education and Research.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.visres.2007.11.015](https://doi.org/10.1016/j.visres.2007.11.015).

References

- Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, 52(6), 1106–1129.
- Berthoz, A., Israel, I., Georges-Francois, P., Grasso, R., & Tsuzuku, T. (1995). Spatial memory of body linear displacement: What is being stored? *Science*, 269, 95–98.
- Bigel, M. G., & Ellard, C. G. (2000). The contribution of non visual information to simple place navigation and distance estimation: An examination of path integration. *Canadian Journal of Experimental Psychology*, 54, 172–184.
- Bremmer, F., & Lappe, M. (1999). The use of optical velocities for distance discrimination and reproduction during visually simulated self motion. *Experimental Brain Research*, 127, 33–42.
- Corlett, J. T., Patla, A. E., & Williams, J. G. (1985). Locomotor estimation of distance after visual scanning by children and adults. *Perception*, 14, 257–263.
- Elliott, D. (1987). The influence of walking speed and prior practice on locomotor distance estimation. *Journal of Motor Behavior*, 19, 476–485.
- Foster, K. H., Gaska, J. P., Nagler, M., & Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology*, 365, 331–363.
- Frenz, H., & Lappe, M. (2005). Absolute travel distance from optic flow. *Vision Research*, 45(13), 1679–1692.
- Frenz, H., Bremmer, F., & Lappe, M. (2003). Discrimination of travel distances from situated optic flow. *Vision Research*, 43, 2173–2183.
- Fukushima, S. S., Loomis, J. M., & Da Silva, J. A. (1997). Visual perception of egocentric distance as assessed by triangulation. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 86–100.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Glasauer, S., Schneider, E., Grasso, R., & Ivanenko, Y. P. (2007). Space–time relativity in self-motion reproduction. *Journal of Neurophysiology*, 97, 451–461.
- Glasauer, S., Amorim, M. A., Vitte, E., & Berthoz, A. (1994). Goal-directed linear locomotion in normal and labyrinthine-defective subjects. *Experimental Brain Research*, 98, 323–335.
- Harris, L. R., Jenkin, M., Zikovitz, D., Redlick, F., Jaekl, P., Jasiobedzka, U., Jenkin, H., & Allison, R. (2002). Simulating self motion I: Cues for the perception of motion. *Virtual Reality*, 6(2), 75–85.
- Harris, L., Jenkin, M., & Zikovitz, D. (2000). Visual and non-visual cues in the perception of linear self-motion. *Experimental Brain Research*, 135, 12–21.
- Israel, I., Grasso, R., Georges-Francois, P., Tsuzuku, T., & Berthoz, A. (1997). Spatial memory and path integration studied by self-driven passive linear displacement. *Journal of Neurophysiology*, 77, 3180–3192.
- Lappe, M., Jenkin, M., & Harris, L. R. (2007). Travel distance estimation from visual motion by leaky path integration. *Experimental Brain Research*, 180, 35–48.
- Loomis, J. M., Klatzky, R. L., Philbeck, J. W., & Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Perception & Psychophysics*, 60(6), 966–980.
- Loomis, J. M., Da Silva, J. A., Fujita, N., & Fukushima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 906–921.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- Mittelstaedt, M. L., & Mittelstaedt, H. (2001). Idiothetic navigation in humans: Estimation of path length. *Experimental Brain Research*, 13, 318–332.
- Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: The MIT Press.
- Pavard, B., & Berthoz, A. (1977). Linear acceleration modifies the perceived velocity of a moving visual scene. *Perception*, 6, 529–540.
- Perrone, J. A., & Thiele, A. (2001). Speed skills: Measuring the visual speed analyzing properties of primate MT neurons. *Nature Neuroscience*, 4(5), 526–532.
- Priebe, N. J., Cassanello, C. R., & Lisberger, S. G. (2003). The neural representation of speed in macaque area MT/V5. *Journal of Neuroscience*, 23(13), 5650–5661.
- Perrone, J. A. (2005). Economy of scale: A motion sensor with variable speed tuning. *Journal of Vision*, 26, 28–33, 5(1).
- Reason, J. R. (1978). Motion sickness adaptation: A neural mismatch model. *Journal of the Royal Society of Medicine*, 71(11), 819–829.
- Reason, J. T., & Brand, J. J. (1975). *Motion sickness*. London: Academic Press.
- Redlick, F. P., Jenkin, M., & Harris, L. (2001). Humans can use optic flow to estimate distance of travel. *Vision Research*, 41, 213–219.
- Rieser, J. J., Ashmead, D. H., Talor, C. R., & Youngquist, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception*, 19, 675–689.
- Steenhuis, R. E., & Goodale, M. A. (1988). The effects of time and distance on accuracy of target-directed locomotion: Does an accurate short-term memory for spatial location exist? *Journal of Motor Behavior*, 20, 399–415.
- Sun, H.-J., Campos, J. L., Chan, G. S. W., Young, M., & Ellard, C. (2004). The contributions of static visual cues, non-visual cues, and optic flow in distance estimation. *Perception*, 33, 49–65.
- Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, 22, 377–380.
- Thomson, J. A. (1983). Is continuous visual monitoring necessary in visually guided locomotion? *Journal of Experimental Psychology: Human Perception and Performance*, 9, 427–443.